Generalized Method of Moments

Econometrics I

Ricardo Mora

Department of Economics
Universidad Carlos III de Madrid
Master in Industrial Economics and Markets

Outline

Motivation

The Analogy Principle

- The Generalized Method of Moments (GMM) is a framework for deriving estimators
- GMM estimators use assumptions about the **moments** of the variables to derive an objective function
- The assumed moments of the random variables provide population moment conditions
- We use the data to compute the **analogous sample moment conditions**
- GMM estimates make the sample moment conditions as true as possible: This step is implemented by minimizing an objective function

Method of Moments

Method of Moments

- The Method of Moments (MM) is a particular type of GMM
- In the Method of Moments (MM), we have the same number of sample moment conditions as we have parameters
- In GMM, we may have more sample moment conditions than we have parameters

Unconditional Mean

We want to estimate $\mu = E[y]$

- Population condition: $E[y] - \mu = 0$
- The sample moment condition

$$\frac{1}{N}\sum_{i=1}^{N} y_i - \hat{\mu}^{GMM} = 0$$

- The GMM is obtained by solving the sample moment condition

$$\hat{\mu}^{GMM} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

This method dates back to Pearson (1895)

Ordinary Least Squares

Ordinary least squares (OLS) is an MM estimator

$$y_i = \beta x_i + u_i, \qquad E\left[u_i | x_i\right] = 0$$

- $E\left[y_i - \beta x_i | x_i\right] = 0 \Rightarrow E\left[x_i \left(y_i - \beta x_i\right)\right] = 0$
- $E\left[x_i \left(y_i - \beta x_i\right)\right] = 0$ are the population moment conditions
- The corresponding sample moment conditions:

$$\frac{1}{N}\sum_{i=1}^{N}\left[x_i\left(y_i - \hat{\beta}^{MM} x_i\right)\right] = 0$$

This is the OLS estimator.

Instrumental Variables

Instrumental Variables is an MM estimator

$$y_i = \beta x_i + u_i$$

- $u_i = y_i - \beta x_i$ is such that $E\left[y_i - \beta x_i | z_i\right] = 0$, where $x_i, z_i \in R^k$
- $E\left[y_i - \beta x_i | z_i\right] = 0 \Rightarrow E\left[z_i\left(y_i - \beta x_i\right)\right] = 0$
- $E\left[z_i\left(y_i - \beta x_i\right)\right] = 0$ are the population moment conditions
- The corresponding sample moment conditions:

$$\frac{1}{N}\sum_{i=1}^{N}\left[z_i\left(y_i - \hat{\beta}^{MM} x_i\right)\right] = 0$$

This is the IV estimator.

Generalized Method of Moments

GMM vs. MM

- MM only works when the number of moment conditions equals the number of parameters to estimate
- If there are more moment conditions than parameters, the system of equations is algebraically over-identified and cannot be solved
- GMM chooses the estimates that minimize a quadratic form of the moment conditions
- GMM gets as close to solving the over-identified system as possible
- GMM reduces to MM when the number of parameters equals the number of moment conditions

GMM Definition

- Let θ be a $k \times 1$ vector of parameters such that

$$E\left[m\left(w_i, \theta\right)\right] = 0_{qx1}$$

  where $m()$ is a $q \times 1$, $q \geq k$, vector of known functions and w_i is the data on person i
- For any c, the q sample moments are $m_N(\mathrm{c}) = \frac{1}{N}\sum_{i=1}^{N} m(w_i, c)$

$$\hat{\theta}^{GMM} \equiv \arg\min_{c} m_N(c)' A_N m_N(c)$$

where A_N is a qxq matrix

GMM and the Moments Conditions

$$\hat{\theta}^{GMM} \equiv \arg\min_{c} m_N(c)' A_N m_N(c)$$

where A_N is a qxq matrix

- First Order Conditions:

$$2\nabla m_N A_N m_N \left(\hat{\beta}^{GMM}\right) = 0_k$$

where ∇m_N is a kxq matrix with the k first derivatives of vector m_N

- The GMM estimator imposes in the sample k linear combinations of q moment conditions

Ordinary Least Squares

Ordinary least squares (OLS) is a GMM estimator

$$y_i = \beta x_i + u_i$$

- Define the $kx1$ vector of moments as:

$$m_N(b) = \frac{1}{N}\sum_{i=1}^{N}[x_i(y_i - bx_i)]$$

- $\hat{\beta}^{GMM} \equiv \arg\min_b m_N(b)' A_N m_N(b)$, A_N is kxk
- FOC: $2\nabla m_N A_N m_N\left(\hat{\beta}^{GMM}\right) = 0_k \iff m_N\left(\hat{\beta}^{GMM}\right) = 0_k$
- This is the same system of equations and has the same solution as the OLS estimator.

Instrumental Variables

IV is a GMM estimator

$$y_i = \beta x_i + u_i, \qquad E\left[u_i|z_i\right] = 0$$

- Define the $kx1$ vector of moments as:

$$m_N\left(b\right) = \frac{1}{N}\sum_{i=1}^{N}\left[z_i\left(y_i - bx_i\right)\right]$$

- $\hat{\beta}^{GMM} \equiv \arg\min_b m_N(b)' A_N m_N(b)$, A_N is kxk
- FOC: $2\nabla m_N A_N m_N\left(\hat{\beta}^{GMM}\right) = 0_k \iff m_N\left(\hat{\beta}^{GMM}\right) = 0_k$
- This is the same system of equations and has the same solution as the IV estimator.

2SLS and GMM

2SLS is a GMM estimator for a particular A_N

$$y_i = \beta x_i + u_i, \qquad E\left[u_i | z_i\right] = 0$$

where $x_i \in R^k$ and $z_i \in R^q$, $q > k$

- Take $A_N = \left(\frac{1}{N}\sum_{i=1}^{N}\left(z_i z_i'\right)\right)^{-1}$

The GMM estimator defined by

$$\nabla m_N A_N m_N\left(w_i, \hat{\beta}^{GMM}\right) = 0$$

is the 2SLS estimator with instruments z_i.

Some Properties of GMM

- When $k = q$ (just-identified case), $m_N\left(\hat{\theta}^{MM}\right) = 0 \Rightarrow \hat{\theta}^{MM} \equiv \hat{\theta}^{GMM}$
- Under general conditions, for a given weight matrix A_N, GMM is both consistent and asymptotically normal
  - $\text{plim}\ \hat{\theta}^{GMM} = \theta$
  - $\sqrt{N}\left(\hat{\theta}^{GMM} - \theta\right) \xrightarrow{d} \mathcal{N}(0, W)$ where W can be consistently estimated

Consistency

- Under general conditions, if

1. $\text{plim } A_N = A$ positive definite
2. $\text{plim } m_N(c) = E[m(w, c)]$, and
3. $AE[m(w, c)] = 0$,

then

$$\text{plim } \hat{\theta}^{GMM} = \theta$$

Asymptotic Normality

- Under general conditions, if

1. $\hat{\theta}^{GMM}$ is consistent,
2. $\text{plim}\, \frac{\partial m_N}{\partial c'} = D(c)$,
3. $D'AD$ $(D \equiv D(\theta))$ is non-singular, and
4. $\sqrt{N} m_N(\theta) \xrightarrow{d} \mathcal{N}(0, V)$

then

$$\sqrt{N}\left(\hat{\theta}^{GMM} - \theta\right) \xrightarrow{d} \mathcal{N}(0, W)$$

where $W = (D'AD)^{-1} D'AVAD(D'AD)^{-1}$

The Weight Matrix A_N

- A_N only affects the efficiency of the GMM estimator
- Setting A_N such that $A = I$ yields consistent, but usually inefficient estimates
- Setting A_N such that $A = k\left[AsyVar\left(m_N\left(\theta\right)\right)\right]^{-1}$ for any $k > 0$ yields an efficient GMM estimator
- Hence, in order to obtain an optimal estimator we need a consistent estimate of $AsyVar\left(m_N\left(\theta\right)\right)$
- This can be done in a two-step procedure

An Efficient GMM

- We can take two steps to get an efficient GMM estimator
  1. Get $\hat{\theta}^{GMM1} \equiv \arg\min_c m_N(c)' m_N(c)$ to get $A_N = \left[\hat{Var}\left(m_N\left(\hat{\theta}^{GMM1}\right)\right)\right]^{-1}$
  2. Get $\hat{\theta}^{GMM2} \equiv \arg\min_c m_N(c)' A_N m_N(c)$

2SLS and Efficiency

- Under conditional homoskedasticity, i.e., $Var(u_i|z_i) = \sigma^2 E(z_i z_i')$, 2SLS is optimal
- If conditional homoskedasticity is violated, then 2SLS is not optimal and the standard formula for the asymptotic variance will not be consistent
  - We can still use 2SLS as a consistent estimator
  - To do the correct inference, we would need an estimate of the variance robust to heteroskedasticity
- Under heteroskedasticity, we can estimate a two-step optimal GMM estimator:
  1. Compute 2SLS and obtain $A_N = \left(\frac{1}{N}\sum_{i=1}^{N} \hat{u}_i^2 (z_i z_i')\right)^{-1}$
  2. Obtain the GMM estimator such that $\nabla m_N A_N m_N \left(w_i, \hat{\beta}^{GMM}\right) = 0$

Testing Overidentifying Restrictions

The Sargan test

If all $q > k$ restrictions are right,

$$N m_N(\hat{\theta})' \left[\hat{Var}\left(m_N\left(\hat{\theta}\right)\right)\right]^{-1} m_N(\hat{\theta}) \xrightarrow{d} \chi^2_{q-k}$$

where $\hat{\theta}$ is the most efficient GMM estimator and $\hat{Var}\left(m_N\left(\hat{\theta}\right)\right)$ is a consistent estimator of $Var\left(m_N\left(\hat{\theta}\right)\right)$.

- the basic idea is that when all restrictions are right, then there will be $q-k$ linear combinations of $m_N\left(\hat{\theta}\right)$ that should be close to zero but are not exactly close to zero

Incremental Sargan test

- Suppose we have $q > k$ restrictions.
- We can first test use $q_1 > k$, with $q_1 < q$, restrictions and get $\hat{\theta}^1$ and use the Sargan test

$$S^1 \equiv N m_N\left(\hat{\theta}^1\right)'\left[\hat{Var}\left(m_N\left(\hat{\theta}^1\right)\right)\right]^{-1} m_N\left(\hat{\theta}^1\right) \xrightarrow{d} \chi^2_{q_1-k}$$

- We can also use all $q > k$, restrictions and get $\hat{\theta}$ and use the Sargan test

$$S \equiv N m_N\left(\hat{\theta}\right)'\left[\hat{Var}\left(m_N\left(\hat{\theta}\right)\right)\right]^{-1} m_N\left(\hat{\theta}\right) \xrightarrow{d} \chi^2_{q-k}$$

- We can test the validity of the $q - q_1$ additional restrictions

$$S^d \equiv S - S^1 \xrightarrow{d} \chi^2_{q-q_1}$$

Summary

- OLS, IV, and 2SLS are shown to be particular cases of GMM
- They will be efficient only under restrictive assumptions
- It is always possible to obtain a two-step consistent and asymptotically efficient GMM estimator
- When the model is over-identified, we can test the over-identifying restrictions or a subset of them